06-22-00

A

06/20/00 JC853 U.S. PTO

JC490 U.S. PTO 09/597878 06/20/00

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b)

Attorney Docket No. 002717.P007C

First Named Inventor or Application Identifier Haddock et al.

Express Mail Label No. EL591668144US

ADDRESS TO: **Assistant Commissioner for Patents**
Box Patent Application
Washington, D. C. 20231

APPLICATION ELEMENTS
See MPEP chapter 600 concerning utility patent application contents.

1. X Fee Transmittal Form
(Submit an original, and a duplicate for fee processing)

2. X Specification (Total Pages 38)
(preferred arrangement set forth below)
- Descriptive Title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R & D
- Reference to Microfiche Appendix
- Background of the Invention
- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claims
- Abstract of the Disclosure

3. X Drawings(s) (35 USC 113) (Total Sheets 5)

4. X Oath or Declaration/Power of Attorney (Total Pages 5)

a. ___ Newly Executed (Original or Copy)

b. X Copy from a Prior Application (37 CFR 1.63(d))
(for Continuation/Divisional with Box 17 completed) **(Note Box 5 below)**

i. ___ DELETIONS OF INVENTOR(S) Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).

5. X Incorporation By Reference (useable if Box 4b is checked)
The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.

6. ____ Microfiche Computer Program (Appendix)

7. ____ Nucleotide and/or Amino Acid Sequence Submission
(if applicable, all necessary)
a. ___ Computer Readable Copy
b. ___ Paper Copy (identical to computer copy)
c. ___ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ____ Assignment Papers (cover sheet & documents(s))

9. ____ 37 CFR 3.73(b) Statement (where there is an assignee)

10. ____ English Translation Document (if applicable)

11. _X_ a. Information Disclosure Statement (IDS)/PTO-1449

____ b. Copies of IDS Citations

12. _X_ Preliminary Amendment

13. _X_ Return Receipt Postcard (MPEP 503) (Should be specifically itemized)

14. ____ a. Small Entity Statement(s)

____ b. Statement filed in prior application, Status still proper and desired

15. ____ Certified Copy of Priority Document(s) (if foreign priority is claimed)

16. ____ Other: _____

_____

_____

_____

17. **If a CONTINUING APPLICATION,** check appropriate box and supply the requisite information:

_X_ Continuation ____ Divisional ____ Continuation-in-part (CIP)

of prior application No: 09/018,103

18. **Correspondence Address**

____ Customer Number or Bar Code Label _____ (Insert Customer No. or Attach Bar Code Label here)

or

_X_ Correspondence Address Below

NAME Michael A. DeSanctis Reg. No. 39,957 6/20/2000

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

12400 Wilshire Boulevard 7th Floor, Los Angeles, California 90025

(303) 740-1980 Telephone (303) 740-6962 Facsimile

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EL591668144US

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, DC 20231.

June 20, 2000
Date of Deposit

Heather S. South
Name of Person Mailing Correspondence

Signature 6/20/00
Date

Our Docket No: 002717.P007C

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:		
Haddock et al.	Examiner:	Not Yet Assigned
Application No.: Not Yet Assigned	Art Unit:	Not Yet Assigned
Filed: Concurrently Herewith		
For: Policy Based Quality of Service		
This is a Continuation of: Application No.: 09/018,103	Examiner:	Vu, H
Filed: February 3, 1998	Art Unit:	2733

PRELIMINARY AMENDMENT

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Prior to examination of the above-captioned case, the Applicant respectfully requests the Examiner to enter the following amendment and to consider the following remark.

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EL591606814 4US

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, DC 20231.

June 20, 2000
Date of Deposit

Heather S. South
Name of Person Mailing Correspondence

Signature 6/20/00
Date

AMENDMENT

In the Specification:

Please insert as the first sentence of the specification --This is a continuation of application serial no. 09/018,103, filed on February 3, 1998, that is currently pending.--

In the Claims:

For the Examiner's convenience all pending claims are presented herein. Those claims that remain unchanged by this amendment are prefixed with "(Unchanged)". Applicant respectfully requests reconsideration of this application as amended. Please cancel claims 4, 7-8, 21-27, 30-34 without prejudice, and add three new claims (36-38).

Please amend the claims as follows:

1. (Amended) A method [of bandwidth management for use in a packet forwarding device coupled to a network, the method] comprising [the steps of]:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device one or more bandwidth parameters for at least one of the one or more traffic groups[;], the bandwidth parameters including at least a minimum bandwidth parameter indicating a minimum amount of bandwidth the at least one traffic group needs to be provided over a defined time period;

receiving at a first port of a plurality of ports a packet associated with the at least one traffic group; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the one or more bandwidth parameters for the at least one traffic group with which the packet is associated.

2. (Unchanged) The method of claim 1, wherein the network employs a non-deterministic access protocol.

3. (Unchanged) The method of claim 2, wherein the non-deterministic access protocol is Carrier Sense Multiple Access with Collision Detection (CSMA/CD).

5. (Amended) The method of claim 1, wherein the one or more bandwidth parameters include [a maximum bandwidth.] an indication regarding a maximum sustained bandwidth the at least one traffic group can realize over a defined time period.

6. (Amended) The method of claim [1,] 5, wherein the one or more bandwidth parameters include an indication regarding a peak bandwidth representing a bandwidth the at least one traffic group may utilize during a particular time interval in excess of the maximum bandwidth.

9. (Amended) The method of claim 1, further comprising [the steps of]:
classifying the packet as being associated with the at least one traffic group; and
determining a quality of service queue with which the at least one traffic group is associated.

10. (Amended) The method of claim 1, further comprising [the steps of]:
enqueuing the packet onto a queue associated with the traffic group;
determining a current bandwidth metric for the queue; and
dequeuing the packet from the queue if the current bandwidth metric meets a predetermined relationship with the one or more bandwidth parameters.

11. (Amended) The method of claim 10, wherein the current bandwidth metric is evaluated periodically at the expiration of a predetermined time period, and

wherein the step of determining a current bandwidth metric for the queue further comprises [the steps of]:

determining an actual bandwidth for a prior time period;

determining a bandwidth metric for the prior time period; and

combining a portion of the actual bandwidth for the prior time period with a portion of the bandwidth metric for the prior time period to arrive at the current bandwidth metric.

12. (Amended) A method of bandwidth management and traffic prioritization for use in a network of devices, the method comprising [the steps of]:

defining at a packet forwarding device information indicative of one or more traffic groups;

defining at the packet forwarding device information indicative of a quality of service (QoS) policy for [at least one of] the one or more traffic groups, the QoS policy including at least a minimum bandwidth parameter indicating a minimum amount of bandwidth the one or more traffic groups need to be provided over a defined time period;

receiving a packet at a first port of a plurality of ports;

identifying a first traffic group of the one or more traffic groups with which the packet is associated; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the QoS policy for the first traffic group, and wherein the scheduling is independent of end-to-end signaling.

13. (Unchanged) The method of claim 12, wherein the network of devices employs a non-deterministic access protocol.

14. (Unchanged) The method of claim 13, wherein the non-deterministic access protocol is Carrier Sense Multiple Access with Collision Detection (CSMA/CD).

15. (Amended) The method of claim 12, further comprising [the steps of]:
providing a plurality of QoS queues; and
mapping the first traffic group to a first QoS queue of the plurality of QoS queues.

16. (Amended) The method of claim 15, further comprising [the steps of]:
determining a current bandwidth metric for each of the plurality of QoS queues;
dividing the plurality of QoS queues into at least a first group and a second group based upon the current bandwidth metrics and a minimum bandwidth requirement associated with each of the plurality of QoS queues;
if the first group includes at least one QoS queue, then transmitting a packet from the at least one QoS queue; otherwise transmitting a packet from a QoS queue associated with the second group.

17. (Amended) A method [of bandwidth management and traffic prioritization for use in a network of devices, the method] comprising [the steps of]:
receiving at a packet forwarding device information indicative of one or more traffic groups;
receiving at the packet forwarding device information defining a quality of service (QoS) policy for at least one of the one or more traffic groups, the QoS policy including at least a minimum bandwidth[;] <u>indicating a minimum amount of bandwidth the at least one traffic group needs to be provided over a defined time period;</u>
providing a plurality of queues at each of a plurality of output ports;
associating the one or more traffic groups with the plurality of queues based upon the minimum bandwidth; and
scheduling a packet for transmission from one of the plurality of queues onto the network.

18. (Unchanged) The method of claim 17, wherein the information indicative of the one or more traffic groups includes Internet Protocol (IP) subnet membership.

19. (Unchanged) The method of claim 18, wherein the information indicative of the one or more traffic groups includes a media access control (MAC) address.

20. (Unchanged) The method of claim 17, wherein the information indicative of the one or more traffic groups includes a virtual local area network (VLAN) identifier.

28. (Amended) A method of bandwidth management for use in a packet forwarding device participating in a connectionless network, the method comprising [the steps of]:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device one or more bandwidth parameters for at least one of the one or more traffic groups[;], the bandwidth parameters including at least a minimum bandwidth indicating a minimum amount of bandwidth the at least one traffic group needs to be provided over a defined time period;

receiving at a first port of a plurality of ports a packet associated with the at least one traffic group; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the one or more bandwidth parameters for the traffic group with which the packet is associated.

29. (Amended) A packet forwarding device for use in a network employing a non-deterministic assess protocol, the packet forwarding device comprising:

an input unit configured to receive information defining one or more traffic groups and one or more associated bandwidth parameters[;], the one or more associated bandwidth parameters including minimum bandwidth indicating a minimum amount of bandwidth the traffic group needs to be provided over a defined time period;

a plurality of ports configured to transmit packets onto an attached network segment, each port having a plurality of queues and configured to select a queue of the plurality of queue from which to transmit a next packet based upon the one or more bandwidth parameters; and

a filtering and forwarding engine coupled to the plurality of ports and configured to process received packets, the filtering and forwarding engine identifying a traffic group of the one or more traffic groups with which a received packet is associated and queuing the received packet for transmission from one of the plurality of ports based upon the identified traffic group.

35. (Amended) A machine-readable medium having stored thereon data representing sequences of instructions, said sequences of instructions which, when executed by a processor, cause said processor to perform [the steps of]:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device information defining a quality of service (QoS) policy for at least one of the one or more traffic groups[;], the QoS policy including at least a minimum bandwidth indicating a

minimum amount of bandwidth the at least one traffic groups needs be provided over a defined time period;

receiving a packet at a first port of a plurality of ports;

identifying a first traffic group of the one or more traffic groups with which the packet is associated; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the QoS policy for the first traffic group, and wherein the scheduling is independent of end-to-end signaling.

Please add the following new claims:

36. (New) The method of claim 1, wherein QoS profile attributes associated with each of the one or more traffic groups include a maximum delay, specifying a time period beyond which further delay cannot be tolerated for the particular traffic group.

37. (New) The method of claim 1, wherein the other QoS profile attributes associated with each of the one or more traffic groups include a Relative Priority, defining the relative importance of a particular traffic group with respect to other traffic groups.

38. (New) A method comprising:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device one or more bandwidth parameters for at least one of the one or more traffic groups;

receiving at a first port of a plurality of ports a packet associated with the at least one traffic group;

enqueuing the packet onto a queue associated with the at least one traffic group;

scheduling the packet for transmission from a second port of the plurality of ports based upon the one or more bandwidth parameters for the at least one traffic group with which the packet is associated by

periodically evaluating a current bandwidth metric for the queue; by

determining an actual bandwidth for a prior time period;

determining a bandwidth metric for the prior time period; and

combining a portion of the actual bandwidth for the prior time period with a portion of the bandwidth metric for the prior time period to arrive at the current bandwidth metric; and

dequeuing the packet from the queue if the current bandwidth metric meets a predetermined relationship with the one or more bandwidth parameters.

REMARK

Applicant respectfully requests reconsideration of this application as amended. Claims 1, 5-6, 9-12, 15-17, 28, 29 and 35 have been amended. Claims 4, 7-8, 21-27, 30-34 have been cancelled without prejudice, and three new claims (36-38) have been added. Therefore, claims 1-3, 5-6, 9-20, 28-29 and 35-38 are presented for examination.

35 U.S.C. §112 rejection,

second paragraph

In the Patent case, the Examiner rejected claims 24-26 and 34 under 35 U.S.C. §112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

The Applicant submits herein proposed amendments, which are thought to overcome the reasons for rejection. Accordingly, the Applicant respectfully requests that the rejection be withdrawn.

35 U.S.C. §102 Rejection

In the Patent case, the Examiner relied upon U.S. Patent No. 5,499,238 of Shon and U.S. Patent No. 5,381,413 Tobagi et al. (hereinafter "Tobagi"). The Applicant respectfully submits the claims, as amended, are thought to overcome the reasons for rejection. The independent claims 1, 12, 17, 28, 29 and 35 have been amended to include a limitation thought to distinguish the present invention from the references relied on by the Examiner. Accordingly, the Applicant respectfully requests that this rejection be withdrawn. The Applicant respectfully submits the following arguments pointing out

significant differences between independent claims 1, 12, 17, 28, 29 and 35 submitted by the Applicant and Shon and Tobagi.

Shon

As amended, in a pertinent part, claim 1 recites "the bandwidth parameters including at least a minimum bandwidth parameter indicating a minimum amount of bandwidth the at least one traffic group needs to be provided over a defined time period." As understood by the Applicant, Shon grants priority if the buffer is full (see col. 7, lines 13-16), hence relying on relative priority to define the importance of a particular traffic group, where traffic groups with a higher priority are preferred over those with lower priorities. The Applicant's invention as recited in claim 1, however, requires the use of bandwidth parameters including at least an indication regarding a minimum amount of bandwidth a particular traffic group needs to be provided over a defined period. Importantly, if the sum of the minimum bandwidths for all traffic groups defined is less than or equal to 100% of the available bandwidth, then the scheduling processing can assure that each traffic group will receive at least the minimum bandwidth requested. The set of bandwidth parameters, including minimum, maximum and peak bandwidths, in combination with the variety of traffic classification schemes gives a network manager enormous control and flexibility in prioritizing and managing traffic flowing through packet forwarding devices in a network, versus relying on relative priority (see specification, pages 3, 16 and 19-20).

As understood by the Applicant, Shon does not teach the use of bandwidth parameters of the present invention and instead relies on the method of relative priority. The absence of the element significantly limits Shon, while it plays an important role in the invention claimed by the Applicant. Accordingly, for at least this reason, the Applicant respectfully requests the withdrawal of the rejection to this claim.

With regard to independent claims 12, 17, 28 and 35, as amended, limitations similar to those discussed with respect to claim 1 apply. Therefore, the Applicant respectfully requests the withdrawal of the rejection to these claims.

Tobagi

As amended, in a pertinent part, claim 29 recites "an input unit configured to receive information defining one or more traffic groups and one or more associated bandwidth parameters, the one or more associated bandwidth parameters including minimum bandwidth indicating a minimum amount of bandwidth the traffic group needs to be provided over a defined time period." As understood by the Applicant, Tobagi does not specifically teach configuration of the input unit to receive one or more bandwidth parameters as claimed by the Applicant. Additionally, Tobagi does not specifically disclose selecting a queue of the plurality of queues from which to transmit a next packet based upon at least a minimum queue bandwidth requirement. As understood by the Applicant, Tobagi employs a conventional station, where packets of all types are submitted to the MAC layer, which transmits them in an arbitrary order (see col. 7, lines 52-54), where each station, in Tobagi, is provided with a throttler, which is assigned to the function of submitting packets (see col. 7 lines 63-68 and col. 8, line 1). In case of a

backlog, continuous attempts are made to transmit packets of each type whenever possible (see col. 7, lines 54-58). For at least this reason, claim 29 is distinguishable over Tobagi.

With regard to independent claim 35, as amended, limitations similar to those discussed with respect to claim 29 apply. Accordingly, for at least this reason, the Applicant respectfully requests the withdrawal of the rejection to this claim.

35 U.S.C. §103 Rejection,

Shon in view of Tobagi et al.

In the parent case, the Examiner rejected claims 3, 14 and 19 under 35 U.S.C. §103(a) as being unpatentable over US Patent No. 5,499,238 of Shon, in view of US Patent No. 5,381,413 of Tobagi. The Applicant respectfully disagrees with the Examiner's characterization of this combination of references. The Applicant respectfully submits the claims, as amended, are thought to overcome the reasons for rejection. Claims 3, 14 and 19 are dependent claims including the limitations of the, amended, independent claims as discussed above. Claim 3 depends on claim 2, which depends on independent claim 1. Claim 14 depends on claim13, which depends on independent claim 12. Claim 19 depends on claim 18, which depends on independent claim 17.

Again, as understood by the Applicant, Shon generally relates to a relative priority scheme. Shon grants priority if a buffer is full (see col. 7, lines 13-16), in order to define the importance of a particular traffic group. As discussed above, each of the independent claims recite the use of bandwidth parameters indicating at least a minimum amount of

bandwidth a particular traffic group needs to be provided over a defined period. Briefly, as understood by the Applicant, neither Tobagi nor Shon, specifically disclose the use of the recited bandwidth parameters. For at least this reason and the reasons discussed above with reference to the independent claims, the claims are thought to be allowable over the combination of Shon and Tobagi.

Conclusion

For the reasons cited above, claims 1-3, 5-6, 9-20, 28-29 and 35-38 are thought to be in condition for allowance. If the Examiner finds any remaining impediment to the prompt allowance of these claims that could be clarified with a telephone conference, the Examiner is respectfully requested to contact Michael DeSanctis at (303) 740-1980.

Charge our Deposit Account

Please charge any shortage to our Deposit Account No. 02-2666.

Respectfully submitted,

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

Date: 6/20/2000

Michael Anthony DeSanctis
Reg. No. 39,957

12400 Wilshire Boulevard
7th Floor
Los Angeles, California 90025-1026
(303) 740-1980

Attorney Docket No.: 02717.P007 Patent Application

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

TITLE OF THE INVENTION

POLICY BASED QUALITY OF SERVICE

INVENTORS

STEPHEN R. HADDOCK
JUSTIN N. CHUEH
SHEHZAD T. MERCHANT
ANDREW H. SMITH
MICHAEL YIP

Prepared by

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1026
(408) 720-8598

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EL 591606814445

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, DC 20231.

June 20, 2000
Date of Deposit

Heather S South
Name of Person Mailing Correspondence

Signature 6/20/00
Date

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EM492087050US

Date of Deposit: February 3, 1998

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Commissioner of Patents and Trademarks, Washington, D. C. 20231

EDITH FUENTES
(Typed or printed name of person mailing paper or fee)

(Signature of person mailing paper or fee)

2-3-98
(Date signed)

This application claims the benefit of U.S. Provisional Application No. 60/057,371, filed 8/29/97.

COPYRIGHT NOTICE

BACKGROUND OF THE INVENTION

Field of the Invention

The invention relates generally to the field of computer networking devices. More particularly, the invention relates to a flexible, policy-based mechanism for managing, monitoring, and prioritizing traffic within a network and allocating bandwidth to achieve true Quality of Service (QoS).

Description of the Related Art

Network traffic today is more diverse and bandwidth-intensive than ever before. Today's intranets are expected to support interactive multimedia, full-motion video, rich graphic images and digital photography. Expectations about the quality and timely presentation of information received from networks is higher than ever. Increased network speed and bandwidth alone will not satisfy the high demands of today's intranets.

The Internet Engineering Task Force (IETF) is working on a draft standard for the Resource Reservation Protocol (RSVP), an Internet Protocol- (IP) based protocol that

allows end-stations, such as desktop computers, to request and reserve resources within and across networks. Essentially, RSVP is an end-to-end protocol that defines a means of communicating the desired Quality of Service between routers. RSVP is receiver initiated. The end-station that is receiving the data stream communicates its requirements to an adjacent router and those requirements are passed back to all intervening routers between the receiving end-station and the source of the data stream and finally to the source of the data stream itself. Therefore, it should be apparent that RSVP must be implemented across the whole network. That is, both end-stations (e.g., the source and destination of the data stream) and every router in between should be RSVP compliant in order to accommodate the receiving end-station's request.

While RSVP allows applications to obtain some degree of guaranteed performance, it is a first-come, first-served protocol, which means if there are no other controls within the network, an application using RSVP may reserve and consume resources that could be needed or more effectively utilized by some other mission-critical application. A further limitation of this approach to resource allocation is the fact that end-stations and routers must be altered to be RSVP compliant. Finally, RSVP lacks adequate policy mechanisms for allowing differentiation between various traffic flows. It should be appreciated that without a policy system in place, the network manager loses control.

Recent attempts to facilitate traffic differentiation and prioritization include draft standards specified by the Institute of Electrical and Electronics Engineers (IEEE). The IEEE 802.1Q draft standard provides a packet format for an application to specify which Virtual Local Area Network (VLAN) a packet belongs to and the priority of the packet. The IEEE 802.1p committee provides a guideline to classify traffic based on a priority indicator in an 802.1Q frame tag. This allows VLANs to be grouped into eight different

traffic classes or priorities. The IEEE 802.1p committee does not, however, define the mechanism to service these traffic classes.

What is needed is a way to provide true Quality of Service ("QoS") in a network employing a non-deterministic access protocol, such as an Ethernet network, that not only has the ability to prioritize and service different traffic classes, but additionally provides bandwidth management and guarantees a quantifiable measure of service for packets associated with a particular traffic class. More specifically, with respect to bandwidth management, it is desirable to employ a weighted fair queuing delivery schedule which shares available bandwidth so that high priority traffic is usually sent first, but low priority traffic is still guaranteed an acceptable minimum bandwidth allocation. Also, it is desirable to centralize the control over bandwidth allocation and traffic priority to allow for QoS without having to upgrade or alter end-stations and existing routers as is typically required by end-to-end protocol solutions. Further, it would be advantageous to put the control in the hands of network managers by performing bandwidth allocation and traffic prioritization based upon a set of manager-defined administrative policies. Finally, since there are many levels of control a network manager may elect to administer, it is desirable to provide a variety of scheduling mechanisms based upon a core set of QoS profile attributes.

BRIEF SUMMARY OF THE INVENTION

A flexible, policy-based, mechanism for managing, monitoring, and prioritizing traffic within a network and allocating bandwidth to achieve true Quality of Service (QoS) is described. According to one aspect of the present invention, a method is provided for managing bandwidth allocation in a network that employs a non-deterministic access protocol. A packet forwarding device receives information indicative of a set of traffic groups. The packet forwarding device additionally receives parameters, such as bandwidth and priority parameters, corresponding to the traffic groups. After receiving a packet associated with one of the traffic groups on a first port, the packet forwarding device schedules the packet for transmission from a second port based upon parameters corresponding to the traffic group with which the packet is associated. Advantageously, in this manner, a weighted fair queuing schedule that shares bandwidth according to some set of rules may be achieved.

According to another aspect of the present invention, a method is provided for managing bandwidth allocation and traffic prioritization in a packet forwarding device. The packet forwarding device receives information indicative of a set of traffic groups. The packet forwarding device additionally receives information defining a Quality of Service (QoS) policy for the traffic groups. After a packet is received by the packet forwarding device, a traffic group with which the packet is associated is identified. Subsequently, rather than relying on an end-to-end signaling protocol for scheduling, the packet is scheduled for transmission based upon the QoS policy for the identified traffic group. Therefore, bandwidth allocation and traffic prioritization are based upon a set of administrative policies over which the network manager retains control.

According to yet another aspect of the present invention, a number of QoS queues are provided at each port of the packet forwarding device. A current bandwidth metric is determined for each of the QoS queues for a particular port. The QoS queues are divided into two groups based upon their respective bandwidth metrics and their respective minimum bandwidth requirements. Subsequently, the groups are used as a first level arbitration mechanism to select a QoS queue that will source the next packet.

Other features of the present invention will be apparent from the accompanying drawings and from the detailed description which follows.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

Figure 1A is a simplified block diagram of an exemplary switch architecture in which one embodiment of the present invention may be implemented.

Figure 1B is a logical view of the interaction between switch processing blocks according to one embodiment of the present invention.

Figure 2 is a flow diagram illustrating high level bandwidth management and traffic prioritization processing according to one embodiment of the present invention.

Figure 3 is a flow diagram illustrating periodic evaluation of QoS categories according to one embodiment of the present invention.

Figure 4 is a flow diagram illustrating next packet scheduling according to one embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

A flexible, policy-based, mechanism for managing, monitoring, and prioritizing traffic within a network and allocating bandwidth to achieve true Quality of Service (QoS) is described. "Quality of Service" in this context essentially means that there is a quantifiable measure of the service being provided. The measure of service being provided may be in terms of a packet loss rate, a maximum delay, a committed minimum bandwidth, or a limited maximum bandwidth, for example.

In the present invention, a number of QoS queues may be provided at each port of a packet forwarding device, such as a Local Area Network (LAN) switch. Based upon a set of QoS parameters, various types of traffic can be distinguished and associated with particular QoS queues. For example, packets associated with a first traffic group may be placed onto a first QoS queue and packets associated with another traffic group may be placed onto a second QoS queue. When a port is ready to transmit the next packet, a scheduling mechanism may be employed to select which QoS queue of the QoS queues associated with the port will provide the next packet for transmission.

In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art that the present invention may be practiced without some of these specific details. In other instances, well-known structures and devices are shown in block diagram form.

The present invention includes various steps, which will be described below. The steps of the present invention may be performed by hardware components or may be embodied in machine-executable instructions, which may be used to cause a general-purpose or special-purpose processor programmed with the instructions to perform the

steps. Alternatively, the steps may be performed by a combination of hardware and software. While, embodiments of the present invention will be described with reference to a high speed Ethernet switch, the method and apparatus described herein are equally applicable to other types of network devices or packet forwarding devices.

An Exemplary Switch Architecture

An overview of the architecture of a switch 100 in which one embodiment of the present invention may be implemented is illustrated by Figure 1A. The central memory architecture depicted includes multiple ports 105 and 110 each coupled via a channel to a filtering/forwarding engine 115. Also coupled to the filtering/forwarding engine 115 is a forwarding database 120, a packet Random Access Memory (RAM) 125, and a Central Processing Unit (CPU) 130.

According to one embodiment, each channel is capable of supporting a data transfer rate of one gigabit per second in the transmit direction and one gigabit per second in the receive direction, thereby providing 2 Gb/s full-duplex capability per channel. Additionally, the channels may be configured to support one Gigabit Ethernet network connection or eight Fast Ethernet network connections.

The filtering/forwarding engine 115 includes an address filter (not shown), a switch matrix (not shown), and a buffer manager (not shown). The address filter may provide bridging, routing, Virtual Local Area Network (VLAN) tagging functions, and traffic classification. The switch matrix connects each channel to a central memory such as packet RAM 125. The buffer manager controls data buffers and packet queue structures and controls and coordinates accesses to and from the packet RAM 125.

The forwarding database 120 may store information useful for making forwarding decisions, such as layer 2 (e.g., Media Access Control (MAC) layer), layer 3 (e.g.,

Network layer), and/or layer 4 (e.g., Transport layer) forwarding information, among other things. The switch 100 forwards a packet received at an input port to an output port by performing a search on the forwarding database using address information contained within the header of the received packet. If a matching entry is found, a forwarding decision is constructed that indicates to which output port the received packet should be forwarded, if any. Otherwise, the packet is forwarded to the CPU 130 for assistance in constructing a forwarding decision.

The packet RAM 125 provides buffering for packets and acts as an elasticity buffer for adapting between incoming and outgoing bandwidth differences. Packet buffering is discussed further below.

Logical View of Exemplary Switch Processing

Figure 1B is a logical view of the interaction between exemplary switch processing blocks that may be distributed throughout the switch 100. For example, some of the processing may be performed by functional units within the ports of the switch and other processing may be performed by the CPU 130 or by the address filter/switch matrix/buffer manager 115. In any event, the processing can be conceptually divided into a first group of functions 160 dedicated to input processing and a second group of functions 185 dedicated to output processing. According to the present embodiment, the first group 160 includes a comparison engine 155, an enqueue block 161, a packet classification block 150, and a buffer manager 165. The second group 185 includes a dequeue block 162, a Quality of Service (QoS) category evaluation block 175, and a scheduler 170.

Additionally, a user interface (UI) 145 may be provided for receiving various parameters from the network manager. The UI may be text based or graphical. In one embodiment, the UI 145 may include an in-band HyperText Markup Language (HTML)

browser-based management tool which may be accessed by any standard web browser. In any event, the goal of the UI 145 is to separate high-level policy components, such as traffic grouping and QoS profiles from the details of the internal switch hardware. Thus, user configuration time is minimized and a consistent interface is provided to the user.

The UI 145 receives information indicative of one or more traffic groups. This information may be provided by the network manager. There are several ways to define a traffic group. Table 1 below illustrates a variety of traffic classification schemes that may be supported by the UI 145.

Traffic Classification

Policy Based Upon	Traffic Group Definition	OSI Layer
Applications	TCP Session UDP Session RSVP Flow	Transport Layer
Network Layer Topology or Groups of Users	Network Layer Protocol Subnet or IP Address VLAN Identifier	Network Layer
End-Station Applications	MAC Address 802.1p or 802.1Q	Link Layer
Physical Topology	Physical Port	Physical Layer

Table 1

The information used to identify a traffic group typically depends upon what terms the QoS policy is defined. If the QoS policy is based on applications, traffic groups may be differentiated at the Transport layer by Transmission Control Protocol (TCP) session or User Datagram Protocol (UDP) session. For example, the network manager may provide information indicative of TCP source and destination ports and IP source and destination addresses to identify traffic groups. However, if the QoS policy is based upon the Network layer topology or groups of users, traffic group definition may be more

convenient by supplying information regarding the Network layer protocol, such as Internet Protocol (IP) or Internetwork Packet Exchange (IPX), the subnet or IP addresses, or VLAN identifiers. If the QoS policy is defined by end-station applications, then Media Access Control (MAC) addresses, IEEE 802.1p priority indications, or IEEE 802.1Q frames may be employed to identify traffic groups. Finally, if the QoS policy is physical topology based, physical port identifiers may be used to differentiate traffic groups.

It should be noted that Table 1 merely presents an exemplary set of traffic group identification mechanisms. From the examples presented herein, additional, alternative, and equivalent traffic grouping schemes and policy considerations will be apparent to those of ordinary skill in the art. For example, other state information may be useful for purposes of packet classification, such as the history of previous packets, the previous traffic load, the time of day, etc.

It is appreciated that traffic classifications based upon the traffic group definitions listed above may result in overlap. Should the network manager define overlapping traffic groups, the UI 145 may issue an error message and reject the most recent traffic group definition, the UI 145 may issue a warning message to the network manager and allow the more specific traffic group definition to override a conflicting general traffic group definition, or the UI 145 may be configured to respond in another manner.

A number of QoS queues 180 may be provided at each of the ports of a packet forwarding device. In one embodiment, a mapping of traffic groups to QoS queues 180 may be maintained. As traffic groups are provided by the network manager, the UI 145 updates the local mapping of traffic groups to QoS queues 180. This mapping process may be a one-to-one mapping of the traffic groups defined by the network manager to the QoS queues 180 or the mapping process may be more involved. For example, there may be more traffic groups than QoS queues 180, in which case, more than one traffic group will

be mapped to a single QoS queue. Some consolidation rules for combining multiple traffic groups into a single QoS queue will be discussed below.

At any rate, by providing a layer of abstraction in this manner, the network manager need not be burdened with the underlying implementation details, such as the number of QoS queues per port and other queuing parameters. Another advantage achieved by this layer of abstraction between the traffic group definitions and the physical QoS queues is the fact that the UI 145 is now decoupled from the underlying implementation. Therefore, the UI 145 need not be updated if the hardware QoS implementation changes. For example, software providing for traffic group definition need not be changed simply because the number of QoS queues per port provided by the hardware changes.

The input data stream is received by the comparison engine 155 from input switch ports (not shown). Under the direction of the packet classification process 150, the comparison engine 155 determines with which of the previously defined traffic groups a packet in the data stream is associated. The packet classification block 150 may employ the traffic group indications provided by the network manager to provide the comparison engine 155 with information regarding locations and fields to be compared or ignored within the header of a received packet, for example. It should be appreciated if the comparison required for traffic classification is straightforward, such as in a conventional packet forwarding device, then the comparison engine 155 and the packet classification block 150 may be combined.

The packet classification block 150 in conjunction with the UI 145 provide a network manager with a flexible mechanism to control traffic prioritization and bandwidth allocation through the switch 100. Importantly, no end-to-end signaling protocol needs to be implemented by the network devices. For example, the end-station that is to receive the data stream need not reserve bandwidth on each of the intermediate devices between it and

the source of the data stream. Rather, a packet forwarding device employing the present invention can provide some benefit to the network without requiring routers and/or end-stations to do anything in particular to identify traffic. Thus, traffic priority may be enforced by the switch 100 and QoS may be delivered to applications without altering routers or end-stations.

According to one embodiment, the buffer manager 165 participates in policy based QoS by controlling the allocation of buffers within the packet RAM 125. Buffers may be dynamically allocated to QoS queues 180 as needed, within constraints established by QoS profile attributes, which are discussed below. The buffer manager 165 may maintain several programmable variables for each QoS queue. For example, a Minimum Buffer Allocation and a Maximum Queue Depth may be provided for each QoS queue. The Minimum Buffer Allocation essentially reserves some minimum number of buffers in the packet RAM 125 for the QoS queue with which it is associated. The Maximum Queue Depth establishes the maximum number of buffers that can be placed on a given QoS queue. The buffer manager 165 also maintains a Current Queue Depth for each QoS queue to assure the maximum depth is not exceeded. For example, before allowing a buffer to be added to a given QoS queue, the buffer manager 165 may compare the Maximum Queue Depth to the Current Queue Depth to ensure the Maximum Queue Depth is not exceeded.

Variables are also maintained for tracking free buffers in the packet RAM 125. At initialization, a Buffers Free Count contains the total number of buffers available in the packet RAM 125 and a Buffers Reserved Count contains the sum of the minimum buffer allocations for the QoS queues 180. As packets are received they are stored in free buffers, and the Buffers Free Count is decremented by the number of buffers used for such storage. After the appropriate QoS queue has been identified the buffer manager 165 instructs the enqueue block 161 to add the packet to the QoS queue. The enqueue block 161 links the

packet to the identified queue provided that the Current Queue Depth is less than the Maximum Queue Depth and either (1) the Current Queue Depth is less than the Minimum Buffer Allocation or (2) the Buffers Reserved Count is less than the Buffers Free Count. Therefore, if a QoS queue exceeds its reserve of buffers (e.g., Minimum Buffer Allocation), to the extent that additional buffers remain free, the QoS queue may continue to grow. Otherwise, the enqueue block 161 will discard the packet, the buffers are returned to the free pool, and the Buffers Free Count is increased by the number of buffers that would have been consumed by the packet. When a packet is successfully linked to a QoS queue, the Current Queue Depth for that QoS queue is increased by the number of buffers used by the packet. If, prior to the addition of the packet to the queue, the Current Queue Depth was less than the Minimum Buffer Allocation then the Buffers Reserved Count is decreased by the lesser of (1) the number of buffers in the packet or (2) the difference between the Current Queue Depth and the Minimum Buffer Allocation.

The QoS category evaluation process 175 separates the QoS queues into a plurality of categories based upon a set of bandwidth parameters. The scheduler 170 uses the grouping provided by the QoS category evaluation process 175 to select an appropriate QoS queue for sourcing the next packet for a particular port. The evaluation of QoS queue categories may be performed periodically or upon command by the scheduler 170, for example. Periodic evaluation of QoS categories and scheduling is discussed in further detail below.

Responsive to the scheduler 170 the dequeue block 162 retrieves a packet from a specified QoS queue. After the packet has been transmitted, the buffer variables are updated. The Buffers Free Count is increased and the Current Queue Depth is decreased by the number of buffers utilized to store the packet. If the resulting Current Queue Depth is less than the Minimum Buffer Allocation, then the Buffers Reserved Count is increased

by the lesser of the number of buffers utilized to store the packet or the difference between the Current Queue Depth and the Minimum Buffer Allocation.

QoS Profile Attributes

Setting QoS policy is a combination of identifying traffic groups and defining QoS profiles for those traffic groups. According to one embodiment, each individual traffic group may be associated with a QoS profile. However, in alternative embodiments, multiple traffic groups may share a common QoS profile. Having described traffic group classification and identification above, QoS profile attributes (also referred to as parameters) will now be discussed.

Several queuing mechanisms may be implemented using one or more of the following parameters associated with a traffic group: (1) minimum bandwidth, (2) maximum bandwidth, (3) peak bandwidth, (4) maximum delay, and (5) relative priority. In general, the minimum, maximum, and peak bandwidth parameter may be expressed in Mbps, a percentage of total bandwidth, or any other convenient representation.

Minimum bandwidth indicates the minimum amount of bandwidth a particular traffic group needs to be provided over a defined time period. If the sum of the minimum bandwidths for all traffic groups defined is less than 100% of the available bandwidth, then the scheduling processing, discussed below, can assure that each traffic group will receive at least the minimum bandwidth requested.

Maximum bandwidth is the maximum sustained bandwidth the traffic group can realize over a defined time period. In contrast, peak bandwidth represents the bandwidth a traffic group may utilize during a particular time interval in excess of the maximum bandwidth. The peak bandwidth parameter may be used to limit traffic bursts for the traffic group with which it is associated. The peak bandwidth also determines how quickly the

traffic group's current bandwidth will converge to the maximum bandwidth. By providing a peak bandwidth value that is much higher than the maximum bandwidth, if sufficient bandwidth is available, the maximum bandwidth will be achieved relatively quickly. In contrast, a peak bandwidth that is only slightly higher that the maximum bandwidth will cause the convergence to the maximum bandwidth to be more gradual.

Maximum delay specifies a time period beyond which further delay cannot be tolerated for the particular traffic group. Packets comprising the traffic group that are forwarded by the switch 100 are guaranteed not to be delayed by more than the maximum delay specified.

Relative priority defines the relative importance of a particular traffic group with respect to other traffic groups. As will be discussed further below, within the same QoS category, traffic groups with a higher priority are preferred over those with lower priorities.

This small set of parameters in combination with the variety of traffic classification schemes gives a network manager enormous control and flexibility in prioritizing and managing traffic flowing through packet forwarding devices in a network. For example, the QoS profile of a video traffic group, identified by UDP session, might be defined to have a high priority and a minimum bandwidth of 5 Mbps, while the QoS profile of an engineering traffic group, identified by VLAN, may be set to a second priority, a minimum bandwidth of 30 Mbps, a maximum bandwidth of 50 Mbps, and a peak bandwidth of 60 Mbps. Concurrently, the QoS profile of a World Wide Web (WWW) traffic group, identified by protocol (e.g., IP), may be set to have a low priority, a minimum bandwidth of 0 Mbps, a maximum bandwidth of 100%, and a peak bandwidth of 100%.

Consolidation Rules

It was mentioned earlier that multiple traffic groups may be mapped to a single QoS queue. This may be accomplished by maintaining an independent set of variables (e.g., minimum bandwidth, maximum bandwidth, peak bandwidth, maximum delay, and relative priority) for each QoS queue in addition to those already associated with each traffic group and following the general consolidation rules outlined below.

Briefly, when the mapping from traffic groups to QoS queues is one-to-one, the determination of a particular QoS queues' attributes is straightforward. The QoS queue's attributes simply equal the traffic group's attributes. However, when combining multiple traffic groups that do not share a common QoS profile onto a single QoS queue, the following general consolidation rules are suggested: (1) add minimum attributes of the traffic groups being combined to arrive at an appropriate minimum attribute for the target QoS queue (e.g., the QoS queue in which the traffic will be merged), (2) use the largest of maximum attributes to arrive at an appropriate value for a maximum attribute for the target QoS queue, and (3) avoid merging traffic groups that have different relative priorities. This last rule suggests the number of priority levels provided should be less than or equal to the number of QoS queues supported by the implementation to assure traffic groups with different priorities are not combined in the same QoS queue.

Importantly, when a network manager has determined that multiple traffic groups will share a common QoS profile, the consolidation rules need not apply, as the network manager has already, in effect, manually consolidated the parameters.

Bandwidth Management and Traffic Prioritization

Having described an exemplary environment in which one embodiment of the present invention may be implemented, bandwidth management and traffic prioritization will now be described with reference to Figure 2. Figure 2 is a flow diagram illustrating the high level bandwidth management and traffic prioritization processing according to one embodiment of the present invention. In this embodiment, at step 210, a manager-defined QoS policy may be received via the UI 145, for example. The QoS policy is a combination of traffic groups and QoS profile attributes corresponding to those traffic groups.

At step 220, a packet is received by the switch 100. Before the packet can be placed onto a QoS queue for transmission, the traffic group to which the packet belongs is identified at step 230. Typically, information in the packet header, for example, can be compared to the traffic group criteria established by the network manager to identify the traffic group to which the packet belongs. This comparison or matching process may be achieved by programming filters in the switch 100 that allow classification of traffic. According to one embodiment, the packet may be identified using the traffic group definitions listed in Table 1.

At step 250, enqueue processing is performed. The packet is added to the rear of the appropriate QoS queue for the identified traffic group. Importantly, if a maximum delay has been assigned to the traffic group with which the packet is associated, then the packet should either be dropped or transmitted within the period specified. According to one embodiment, this may be accomplished by limiting the depth (also referred to as length) of the corresponding QoS queue. Given the minimum bandwidth of the QoS queue and the maximum delay the traffic group can withstand, a maximum depth for the QoS queue can be calculated. If the QoS queue length remains less than or equal to the

maximum length, then the packet is added to the QoS queue. However, if the QoS queue length would exceed the maximum length by the addition, then the packet is dropped.

At step 260, scheduling is performed. The scheduling/dequeuing processing involves determining the appropriate QoS queue group, selecting the appropriate QoS queue within that QoS queue group, and removing the packet at the front of the selected QoS queue. This selected packet will be the next packet the port transmits. Scheduling will be discussed further below.

Evaluation of QoS Categories

According to one embodiment of the present invention, it is advantageous to divide the QoS queues into at least two categories. The categories may be defined based upon the maximum bandwidth, the minimum bandwidth, the peak bandwidth, and the "current bandwidth." The current bandwidth should not be mistaken for a bandwidth at an instant in time, rather the current bandwidth is a moving average that is updated periodically upon the expiration of a predetermined time period. Empirical data suggests this predetermined time period should be on the order of ten packet times, wherein a packet time is the time required to transmit a packet. However, depending upon the environment and the nature of the traffic, a value in the range of one to one hundred packet times may be more suitable.

The members of the first category ("Category A") are those QoS queues which have a current bandwidth that is below their peak bandwidth and below their minimum bandwidth. Members of the second category ("Category B") include those QoS queues that have a current bandwidth that is greater than or equal to their minimum bandwidth, but less than both their maximum bandwidth and their peak bandwidth. The remaining QoS queues (e.g., those having a current bandwidth that is greater than or equal to either the peak bandwidth or the maximum bandwidth) are ineligible for transmission. These QoS

queues that are ineligible for transmission can be considered a third category ("Category C"). With this overview of QoS categories, an exemplary process for periodic evaluation of QoS categories will now be described.

Figure 3 is a flow diagram illustrating periodic evaluation of QoS categories according to one embodiment of the present invention. In this embodiment, at step 310, processing loops until the predetermined evaluation time period has expired. For example, a test may be performed to determine if the current time is greater than or equal to the last evaluation time plus the predetermined evaluation time interval. Alternatively, the evaluation process may be triggered by an interrupt. In any event, when it is time to evaluate the QoS queue categorization, processing continues with step 320.

It will be appreciated that the time interval chosen for the predetermined evaluation time period should not be too long or too short. If the time interval is too long, one QoS queue might be allowed to monopolize the link until its maximum bandwidth is achieved while other QoS queues remain idle. If the time interval is too short, transmitting a single packet or remaining idle for a single packet time may cause the QoS queue to become a member of a different QoS category (e.g., the single transmission may cause the current bandwidth to exceed the maximum bandwidth or the single idle time may cause the current bandwidth to fall below the minimum bandwidth) because the moving average moves very quickly over short time intervals.

At step 330, the current bandwidth for a particular QoS queue is set to the current bandwidth for that QoS queue as calculated in the previous time interval multiplied by a first weighting factor plus the actual bandwidth that particular QoS queue received in the previous time interval multiplied by a second weighting factor, wherein the weighting factors may be selected to achieve the desired level of responsiveness in the current bandwidth metric. For example, it may be desirable to have the current bandwidth

converge to within a certain percentage of a sustained bandwidth if that bandwidth has been sustained for a certain amount of time. Exemplary weighting factors are in the form (w-1)/w and 1/w, respectively. Using weighting factors of 15/16 for the first weighting factor and a value of 1/16 for the second weighting factor, for example, the current bandwidth will reflect 50% of a step within 13 time intervals, 80% of a step within 27 time intervals, and will be within 2% of the sustained bandwidth in approximately 63 time intervals (assuming a maximum and peak bandwidth of 100%). Alternative ratios and current bandwidth metrics will be apparent to those of ordinary skill in the art.

After the current bandwidth has been evaluated for a QoS queue, at step 340, the QoS queue bandwidth parameters can be compared to the current bandwidth to determine to which QoS category the QoS queue belongs. As described above, if (CURR_BW < PEAK_BW) and (CURR_BW < MIN_BW), then the QoS queue is associated with Category A at step 350. If (CURR_BW ≥ MIN_BW) and ((CURR_BW < MAX_BW) and (CURR_BW < PEAK_BW)), then the QoS queue is associated with Category B at step 360. If (CURR_BW ≥ PEAK_BW) or (CURR_BW ≥ MAX_BW), then the QoS queue is associated with Category C at step 370.

At step 380, if all of the QoS queues have been evaluated, then processing branches to step 310; otherwise, processing continues with step 330.

Scheduling Processing

Briefly, at each port, three levels of arbitration may be employed to select the appropriate QoS queue from which to transmit the next packet. The first level of arbitration selects among the QoS categories. Category A is given priority if any member QoS queues have one or more pending packets. Otherwise, a QoS queue with one or more pending packets of Category B is selected. According to one embodiment, the relative priority

assigned to each QoS queue may be used as a second level of arbitration. In this manner, when multiple QoS queues satisfy the first level arbitration, a higher priority QoS queue is favored over a lower priority QoS queue. Finally, when there is a tie at the second level of arbitration (e.g., two or more QoS queues in the same QoS category have the same relative priority), a round robin or least recently used (LRU) scheme may be employed to select from among the two or more QoS queues until the QoS categories are evaluated.

Assuming a periodic evaluation of QoS categories is being performed, the scheduling processing need not include such evaluation and the scheduling processing may be performed as illustrated by Figure 4, according to one embodiment of the present invention. In the embodiment depicted, at step 410, processing loops until the port associated with the group of QoS queues being evaluated indicates it is ready to receive the next packet for transmission. For example, the port may be polled to determine its transmission status. Alternatively, the scheduling process may be triggered by an interrupt. In any event, when the port is ready for the next packet, processing continues with step 420.

At step 420, a QoS category is selected from which a QoS queue will provide the next packet for transmission. As described above, priority is given to the category containing QoS queues with pending data that are below the peak bandwidth and minimum bandwidth (e.g., Category A). However, if no QoS queues meet this criteria, Category B is selected.

At step 430, if multiple QoS queues are members of the selected QoS category, processing continues with step 440; otherwise, processing branches to step 470.

At step 440, the relative priorities of the QoS queues are used to select among the QoS queues of the selected category that have pending data.

At step 450, if two or more QoS queues have the same priority, then processing continues with step 460. Otherwise, if a QoS queue is found to have the highest relative priority, then processing branches to step 470.

At step 460, the tie is resolved by performing round robin or LRU scheduling. That is, until the QoS categories are evaluated, the QoS queues having the same priority will be rotated through in a predetermined order or scheduled such that the QoS queue that has not provided a packet for transmission recently will be given such an opportunity. After selecting a QoS queue in this manner, processing continues with step 470.

At step 470, a packet is dequeued from the selected QoS queue and the packet is transmitted by the port at step 480. This scheduling process may be repeated by looping back to step 410, as illustrated.

Queuing Schemes

A variety of different queuing mechanisms may be implemented using various combinations of the QoS profile attributes discussed above. Table 2 below illustrates how to achieve exemplary queuing mechanisms and corresponding configurations of the QoS profile attributes.

Queuing Mechanism Configurations

Queuing Mechanism	QoS Profile Attribute Value
Strict Priority Queuing	Minimum Bandwidth = 0 % Maximum Bandwidth = 100 % Peak Bandwidth = 100 % Maximum Delay = N/A Relative Priority = $PRIORITY_i$
Round Robin/ Least Recently Used Queuing	Minimum Bandwidth = 0 % Maximum Bandwidth = 100 % Peak Bandwidth = 100 % Maximum Delay = N/A Relative Priority = <same for all queues>
Weighted Fair Queuing	Minimum Bandwidth = >0 % Maximum Bandwidth = MAX_BW_i Peak Bandwidth = $PEAK\_BW_i$ Maximum Delay = N/A Relative Priority = <same for all queues>

Table 2

$PRIORITY_i$ represents a programmable priority value for a particular QoS queue, i. Similarly, MAX_BW_i and $PEAK\_BW_i$ represent programmable maximum bandwidths and peak bandwidths, respectively, for a particular QoS queue, i.

For a strict priority scheme, each QoS queue's minimum bandwidth is set to zero percent, each QoS queue's maximum bandwidth is set to one hundred percent, and each QoS queue's peak is set to one hundred percent. In this manner, the current bandwidth will never be less than the minimum bandwidth, and the current bandwidth will never exceed either the peak bandwidth or the maximum bandwidth. In this configuration, all QoS queues will be associated with Category B since no QoS queues will satisfy the criteria of either Category A or Category B. Ultimately, by configuring the QoS profile

attributes in this manner, the second level of arbitration (e.g., the relative priority of the QoS queues) determines which QoS queue is to source the next packet.

For a pure round robin or least recently used (LRU) scheme, the QoS profile attributes are as above, but additionally all QoS queue priorities are set to the same value. In this manner, the third level of arbitration determines which QoS queue is to source the next packet.

Finally, weighted fair queuing can be achieved by assigning, at least, a value greater than zero percent to the desired minimum bandwidth. By assigning a value greater than zero to the minimum bandwidth parameter, the particular QoS queue is assured to get at least that amount of bandwidth on average because the QoS queue will be associated with Category A until at least its minimum bandwidth is satisfied. Additionally, different combinations of values may be assigned to the peak and maximum bandwidths to prevent a particular QoS queue from monopolizing the link.

Alternative Embodiments

While evaluation of QoS categories has been described above as occurring periodically, this evaluation may also be triggered by the occurrence of a predetermined event. Alternatively, evaluation of QoS categories may take place as part of the scheduling processing rather than as part of a separate periodic background process.

While a relationship between the number of priority levels and the number of QoS queues has been suggested above, it is appreciated that the number of QoS queues may be determined independently of the number of priority levels. Further, it is appreciated that the number of QoS queues provided at each port may be fixed for every port or alternatively a variable number of QoS queues may be provided for each port.

Finally, in alternative embodiments, weighting factors and ratios other than those suggested herein may be used to adjust the current bandwidth calculation for a particular implementation.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

CLAIMS

What is claimed is:

1. A method of bandwidth management for use in a packet forwarding device coupled to a network, the method comprising the steps of:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device one or more bandwidth parameters for at least one of the one or more traffic groups;

receiving at a first port of a plurality of ports a packet associated with the at least one traffic group; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the one or more bandwidth parameters for the at least one traffic group with which the packet is associated.

2. The method of claim 1, wherein the network employs a non-deterministic access protocol.

3. The method of claim 2, wherein the non-deterministic access protocol is Carrier Sense Multiple Access with Collision Detection (CSMA/CD).

4. The method of claim 1, wherein the one or more bandwidth parameters include a minimum bandwidth.

5. The method of claim 1, wherein the one or more bandwidth parameters include a maximum bandwidth.

6. The method of claim 1, wherein the one or more bandwidth parameters include a peak bandwidth.

7. The method of claim 1, wherein the one or more bandwidth parameters include a maximum delay.

8. The method of claim 1, wherein the one or more bandwidth parameters include a relative priority.

9. The method of claim 1, further comprising the steps of:
classifying the packet as being associated with the at least one traffic group; and
determining a quality of service queue with which the at least one traffic group is associated.

10. The method of claim 1, further comprising the steps of:
enqueuing the packet onto a queue associated with the traffic group;
determining a current bandwidth metric for the queue; and
dequeuing the packet from the queue if the current bandwidth metric meets a predetermined relationship with the one or more bandwidth parameters.

11. The method of claim 10, wherein the current bandwidth metric is evaluated periodically at the expiration of a predetermined time period, and wherein the step of determining a current bandwidth metric for the queue further comprises the steps of:
determining an actual bandwidth for a prior time period;
determining a bandwidth metric for the prior time period; and

combining a portion of the actual bandwidth for the prior time period with a portion of the bandwidth metric for the prior time period to arrive at the current bandwidth metric.

12. A method of bandwidth management and traffic prioritization for use in a network of devices, the method comprising the steps of:

defining at a packet forwarding device information indicative of one or more traffic groups;

defining at the packet forwarding device information indicative of a quality of service (QoS) policy for at least one of the one or more traffic groups;

receiving a packet at a first port of a plurality of ports;

identifying a first traffic group of the one or more traffic groups with which the packet is associated; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the QoS policy for the first traffic group, and wherein the scheduling is independent of end-to-end signaling.

13. The method of claim 12, wherein the network of devices employs a non-deterministic access protocol.

14. The method of claim 13, wherein the non-deterministic access protocol is Carrier Sense Multiple Access with Collision Detection (CSMA/CD).

15. The method of claim 12, further comprising the steps of:

providing a plurality of QoS queues; and

mapping the first traffic group to a first QoS queue of the plurality of QoS queues.

16. The method of claim 15, further comprising the steps of:

determining a current bandwidth metric for each of the plurality of QoS queues;

dividing the plurality of QoS queues into at least a first group and a second group based upon the current bandwidth metrics and a minimum bandwidth requirement associated with each of the plurality of QoS queues;

if the first group includes at least one QoS queue, then transmitting a packet from the at least one QoS queue; otherwise transmitting a packet from a QoS queue associated with the second group.

17. A method of bandwidth management and traffic prioritization for use in a network of devices, the method comprising the steps of:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device information defining a quality of service (QoS) policy for at least one of the one or more traffic groups, the QoS policy including at least a minimum bandwidth;

providing a plurality of queues at each of a plurality of output ports;

associating the one or more traffic groups with the plurality of queues based upon the minimum bandwidth; and

scheduling a packet for transmission from one of the plurality of queues onto the network.

18. The method of claim 17, wherein the information indicative of the one or more traffic groups includes Internet Protocol (IP) subnet membership.

19. The method of claim 18, wherein the information indicative of the one or more traffic groups includes a media access control (MAC) address.

20. The method of claim 17, wherein the information indicative of the one or more traffic groups includes a virtual local area network (VLAN) identifier.

21. A method of bandwidth management and traffic prioritization for use in a network of devices, the method comprising the steps of:

providing a plurality of quality of service (QoS) queues at each of a plurality of output ports, each of the plurality of QoS queues associated with a minimum queue bandwidth requirement;

adding a packet to one of the plurality of QoS queues based upon a traffic group with which the packet is associated; and

scheduling a next packet for transmission onto the network from one of the plurality of QoS queues at a particular output port of the plurality of output ports by:

determining a current bandwidth metric for each of the plurality of QoS queues,

dividing the plurality of QoS queues into at least a first group and a second group based upon the current bandwidth metrics and the minimum queue bandwidth requirements, and

if at least one QoS queue of the plurality of QoS queues, so divided, is associated with the first group, then transmitting a packet from the at least one QoS queue; otherwise transmitting a packet from a QoS queue of the plurality of QoS queues associated with the second group.

22. The method of claim 21, wherein the current bandwidth for a particular QoS queue is calculated as follows:

$CURR\_BW_i = W1 \times CURR\_BW_i + W2 \times ACT\_BW_i$;

where:

$CURR\_BW_i$ represents the current bandwidth for a particular QoS queue,

W1 represents a first weighting factor,

W2 represents a second weighting factor, and

ACT_BW_i represents the actual bandwidth received by the particular QoS queue in a previous time interval.

23. The method of claim 22, wherein W1 = (W-1)/W, W2 = 1/W, and the previous time interval is the most recent time interval.

24. The method of claim 21, further comprising the step of selecting among QoS queues in the same group based upon relative queue priorities associated with the QoS queues.

25. The method of claim 21, further comprising the step of selecting among QoS queues in the same group based upon a round robin selection scheme.

26. The method of claim 21, further comprising the step of selecting among QoS queues in the same group based upon a least recently used (LRU) selection scheme.

27. The method of claim 21, wherein the first group comprises QoS queues associated with a minimum queue bandwidth requirement that is less than the corresponding QoS queue's current bandwidth metric, and wherein the second group comprises

QoS queues associated with a minimum queue bandwidth requirement that is greater than or equal to the corresponding QoS queue's current bandwidth metric.

28. A method of bandwidth management for use in a packet forwarding device participating in a connectionless network, the method comprising the steps of:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device one or more bandwidth parameters for at least one of the one or more traffic groups;

receiving at a first port of a plurality of ports a packet associated with the at least one traffic group; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the one or more bandwidth parameters for the traffic group with which the packet is associated.

29. A packet forwarding device for use in a network employing a non-deterministic assess protocol, the packet forwarding device comprising:

an input unit configured to receive information defining one or more traffic groups and one or more associated bandwidth parameters;

a plurality of ports configured to transmit packets onto an attached network segment, each port having a plurality of queues and configured to select a queue of the plurality of queue from which to transmit a next packet based upon the one or more bandwidth parameters; and

a filtering and forwarding engine coupled to the plurality of ports and configured to process received packets, the filtering and forwarding engine identifying a traffic group of the one or more traffic groups with which a received packet

is associated and queuing the received packet for transmission from one of the plurality of ports based upon the identified traffic group.

30. A packet forwarding device for use in a network employing a non-deterministic assess protocol, the packet forwarding device comprising:

a filtering and forwarding engine configured to forward received packets based upon a traffic group with which the packet is associated; and

a plurality of ports coupled to the filtering and forwarding engine, each port of the plurality of ports configured to receive packets from the filtering and forwarding engine, each port of the plurality of ports having a plurality of Quality of Service (QoS) queues associated with a minimum queue bandwidth requirement, each port of the plurality of ports further configured to schedule a packet for transmission onto the network by

determining a current bandwidth metric for each of the plurality of QoS queues,

dividing the plurality of QoS queues into at least a first group and a second group based upon the current bandwidth metrics and the minimum queue bandwidth requirements, and

if at least one QoS queue of the plurality of QoS queues, so divided, is associated with the first group, then transmitting a packet from the at least one QoS queue; otherwise transmitting a packet from a QoS queue of the plurality of QoS queues associated with the second group.

31. The packet forwarding device of claim 30, wherein the plurality of ports are further configured to select among QoS queues in the same group based upon relative queue priorities associated with the QoS queues.

32. The packet forwarding device of claim 30, wherein the plurality of ports are further configured to select among QoS queues in the same group based upon a round robin selection scheme.

33. The packet forwarding device of claim 30, wherein the plurality of ports are further configured to select among QoS queues in the same group based upon a least recently used (LRU) selection scheme.

34. The packet forwarding device of claim 30, wherein the first group comprises QoS queues associated with a minimum queue bandwidth requirement that is less than the corresponding QoS queue's current bandwidth metric, and wherein the second group comprises QoS queues associated with a minimum queue bandwidth requirement that is greater than or equal to the corresponding QoS queue's current bandwidth metric.

35. A machine-readable medium having stored thereon data representing sequences of instructions, said sequences of instructions which, when executed by a processor, cause said processor to perform the steps of:

receiving at a packet forwarding device information indicative of one or more traffic groups;

receiving at the packet forwarding device information defining a quality of service (QoS) policy for at least one of the one or more traffic groups;

receiving a packet at a first port of a plurality of ports;

identifying a first traffic group of the one or more traffic groups with which the packet is associated; and

scheduling the packet for transmission from a second port of the plurality of ports based upon the QoS policy for the first traffic group, and wherein the scheduling is independent of end-to-end signaling.

ABSTRACT OF THE DISCLOSURE

A flexible, policy-based, mechanism for managing, monitoring, and prioritizing traffic within a network and allocating bandwidth to achieve true quality of service (QoS) is provided. According to one aspect of the present invention, a method is provided for managing bandwidth allocation in a network that employs a non-deterministic access protocol, such as an Ethernet network. A packet forwarding device receives information indicative of a set of traffic groups, such as: a MAC address, or IEEE 802.1p priority indicator or 802.1Q frame tag, if the QoS policy is based upon individual station applications; or a physical port if the QoS policy is based purely upon topology. The packet forwarding device additionally receives bandwidth parameters corresponding to the traffic groups. After receiving a packet associated with one of the traffic groups on a first port, the packet forwarding device schedules the packet for transmission from a second port based upon bandwidth parameters corresponding to the traffic group with which the packet is associated. According to another aspect of the present invention, a method is provided for managing bandwidth allocation in a packet forwarding device. The packet forwarding device receives information indicative of a set of traffic groups. The packet forwarding device additionally receives information defining a QoS policy for the traffic groups. After a packet is received by the packet forwarding device, a traffic group with which the packet is associated is identified. Subsequently, rather than relying on an end-to-end signaling protocol for scheduling, the packet is scheduled for transmission based upon the QoS policy for the identified traffic group.

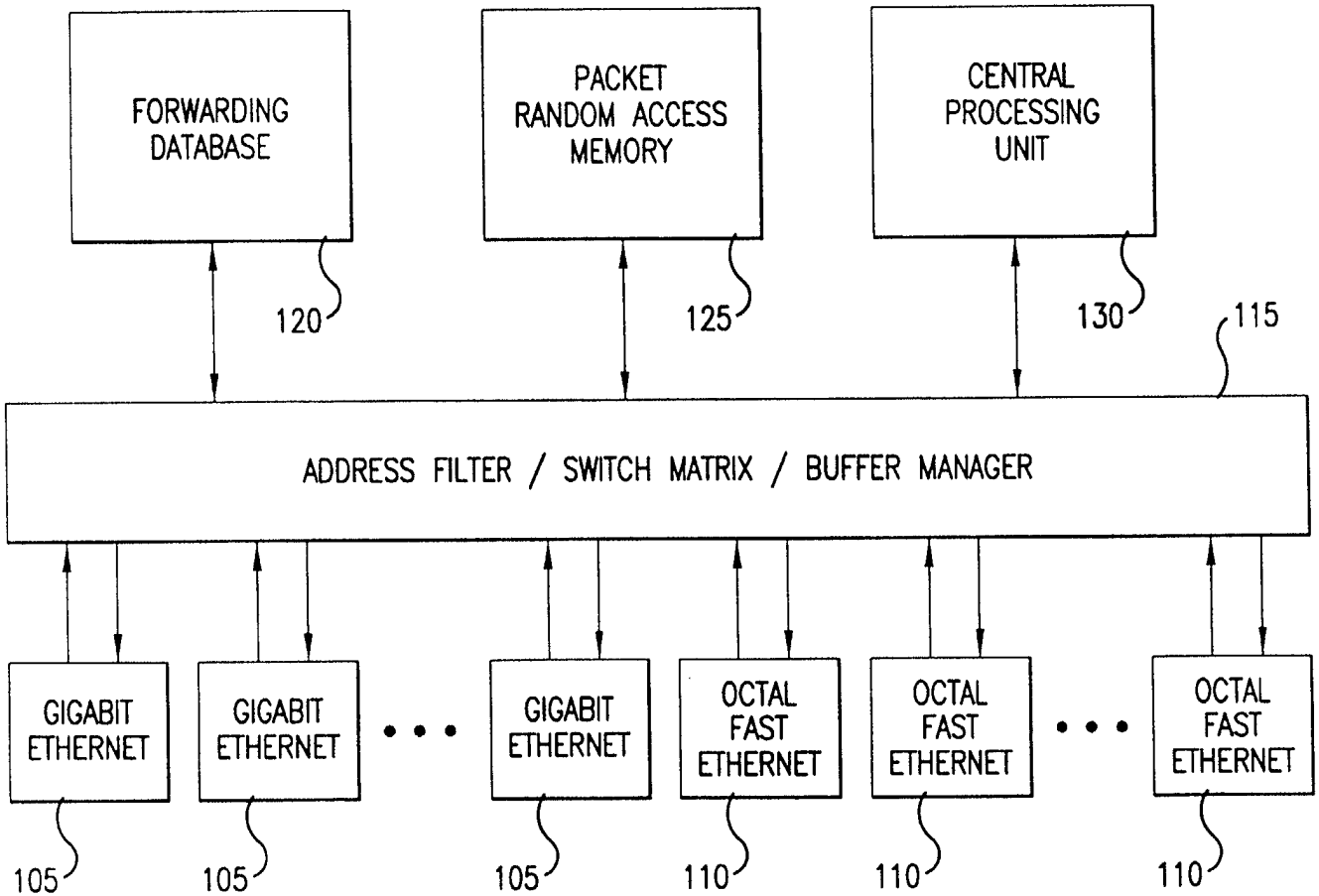
FORWARDING DATABASE
PACKET RANDOM ACCESS MEMORY
CENTRAL PROCESSING UNIT
120
125
130
115
ADDRESS FILTER / SWITCH MATRIX / BUFFER MANAGER
GIGABIT ETHERNET
GIGABIT ETHERNET
GIGABIT ETHERNET
OCTAL FAST ETHERNET
OCTAL FAST ETHERNET
OCTAL FAST ETHERNET
105
105
105
110
110
110
100

FIG.1A

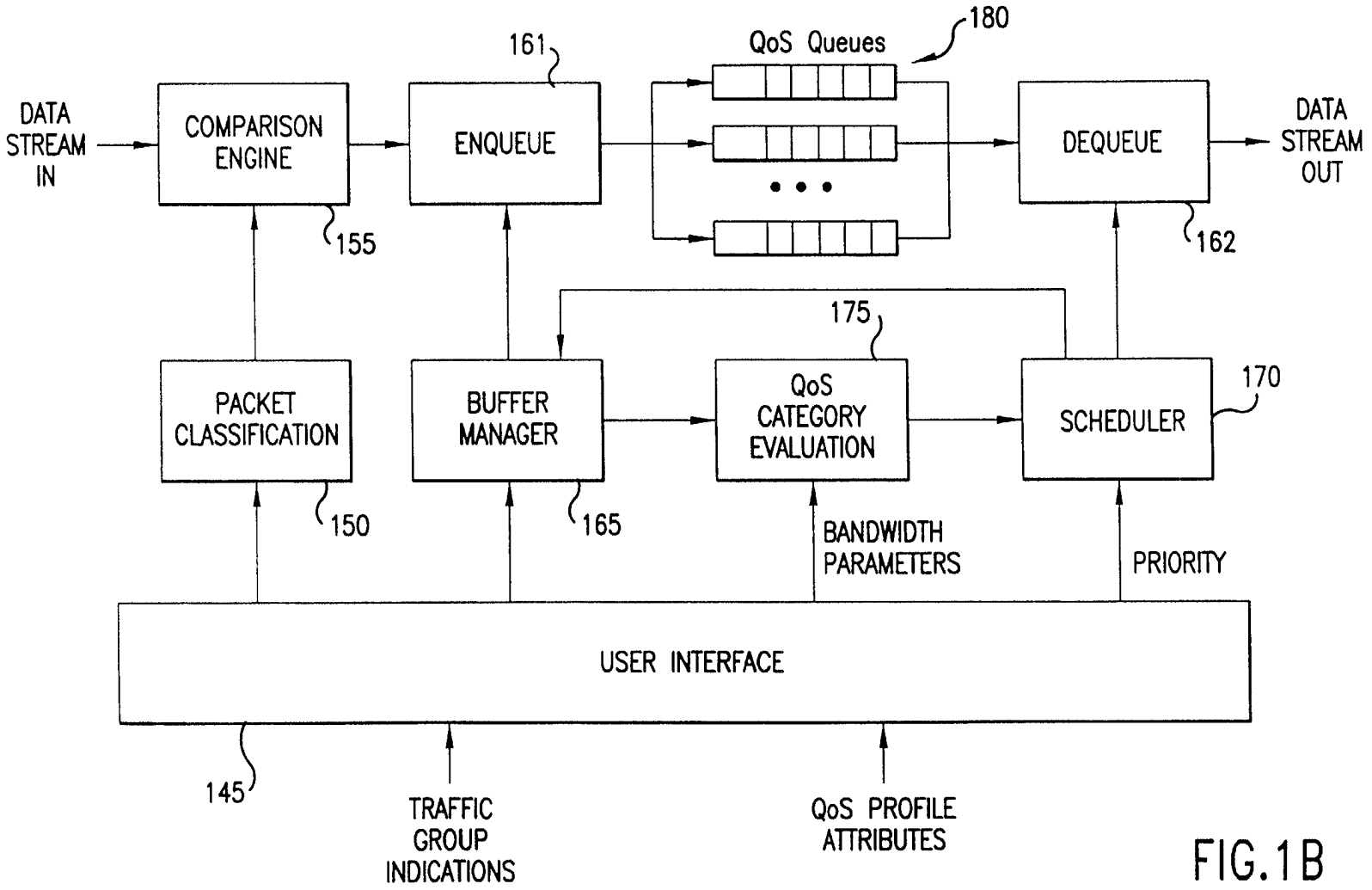

FIG.1B

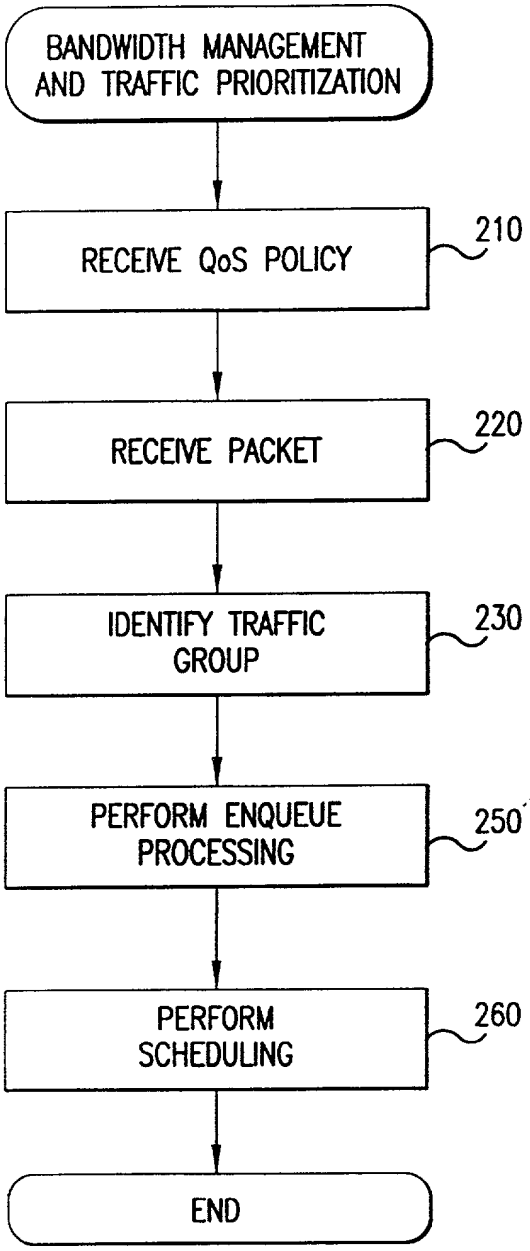
BANDWIDTH MANAGEMENT AND TRAFFIC PRIORITIZATION
RECEIVE QoS POLICY
210
RECEIVE PACKET
220
IDENTIFY TRAFFIC GROUP
230
PERFORM ENQUEUE PROCESSING
250
PERFORM SCHEDULING
260
END

FIG.2

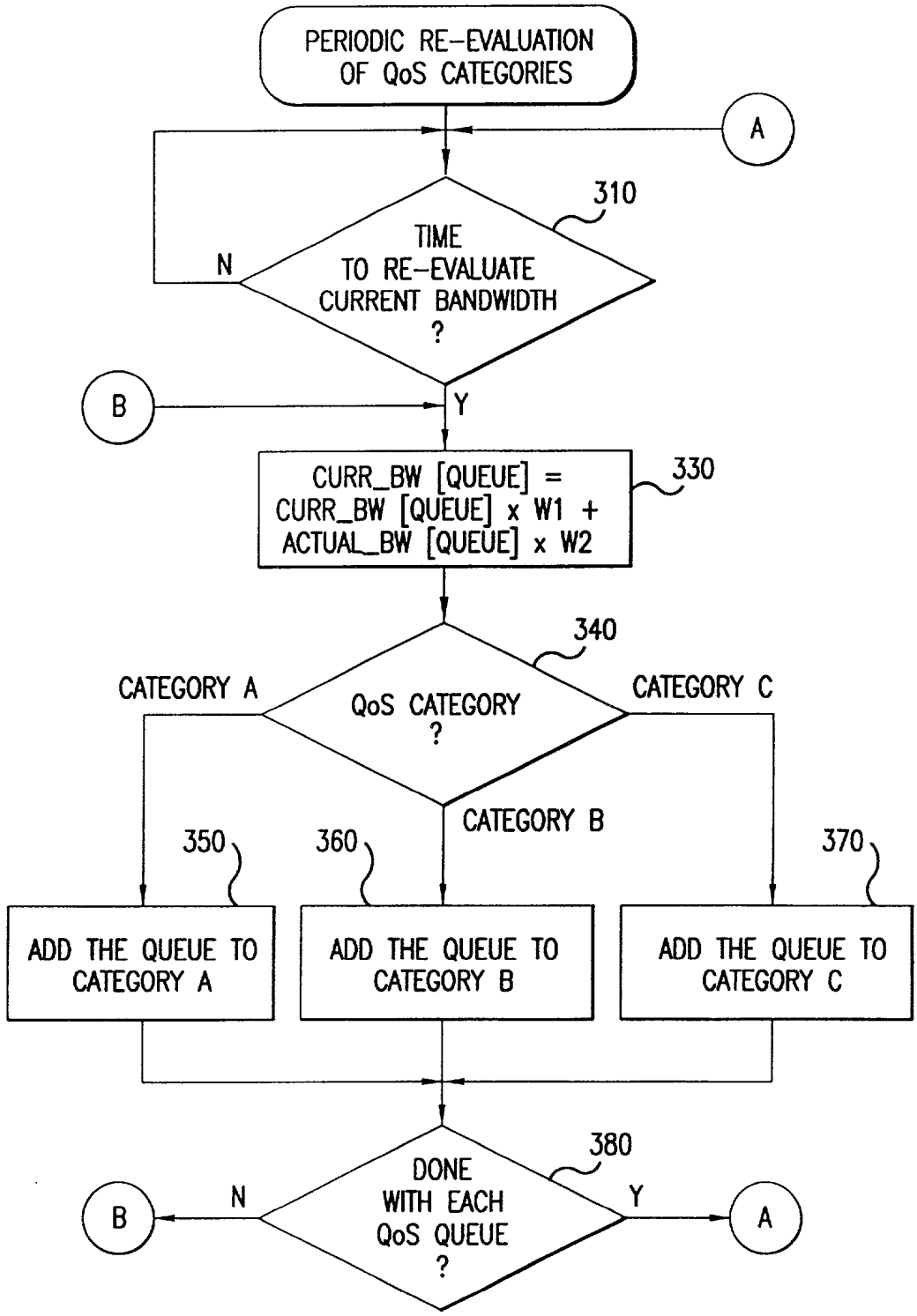
PERIODIC RE-EVALUATION OF QoS CATEGORIES
A
310
TIME TO RE-EVALUATE CURRENT BANDWIDTH ?
N
B
Y
CURR_BW [QUEUE] = CURR_BW [QUEUE] x W1 + ACTUAL_BW [QUEUE] x W2
330
340
CATEGORY A
QoS CATEGORY ?
CATEGORY C
CATEGORY B
350
360
370
ADD THE QUEUE TO CATEGORY A
ADD THE QUEUE TO CATEGORY B
ADD THE QUEUE TO CATEGORY C
380
DONE WITH EACH QoS QUEUE ?
N
Y
B
A

FIG.3

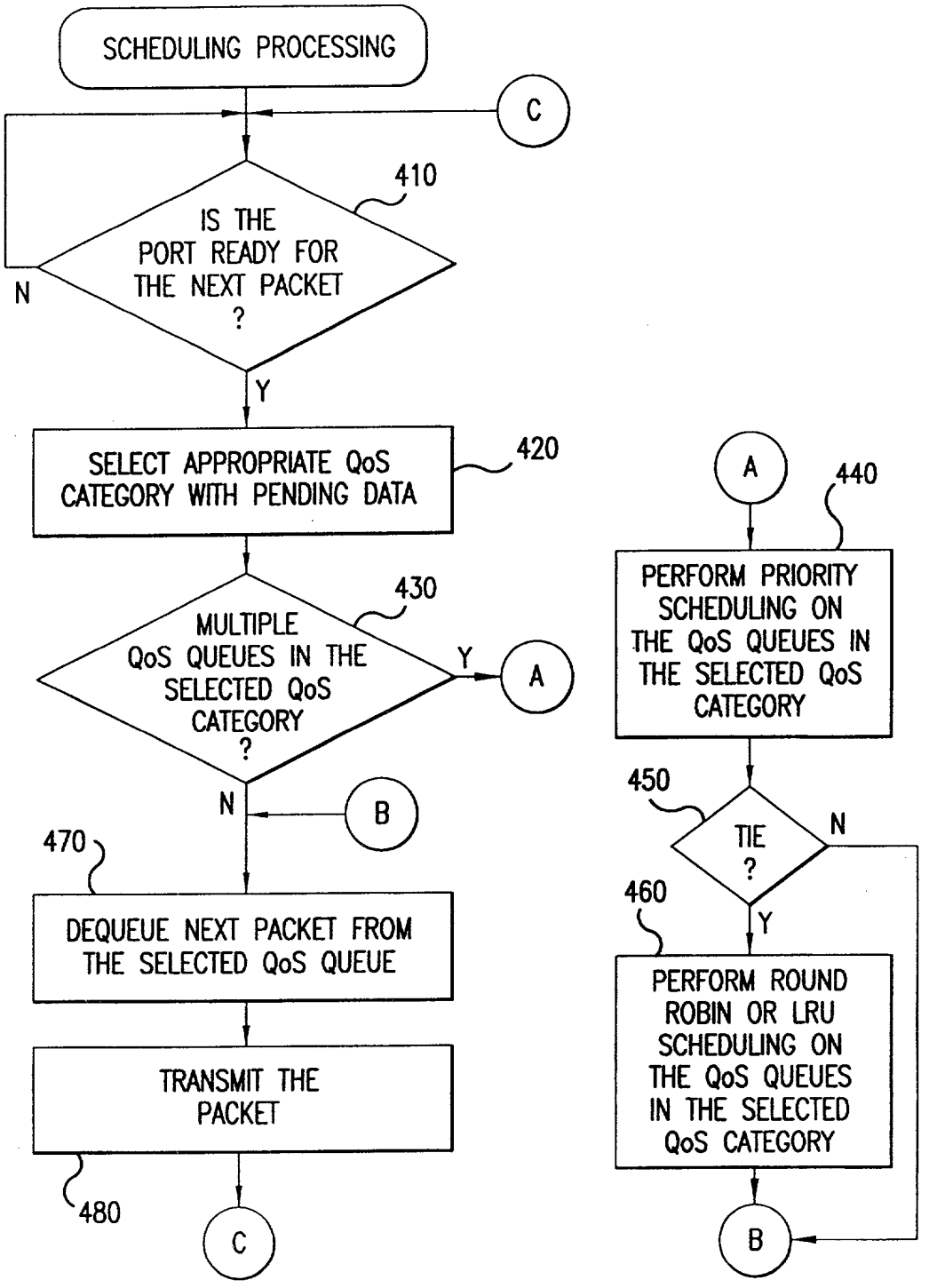
SCHEDULING PROCESSING
C
410
IS THE PORT READY FOR THE NEXT PACKET ?
N
Y
SELECT APPROPRIATE QoS CATEGORY WITH PENDING DATA
420
430
MULTIPLE QoS QUEUES IN THE SELECTED QoS CATEGORY ?
Y
A
N
B
470
DEQUEUE NEXT PACKET FROM THE SELECTED QoS QUEUE
TRANSMIT THE PACKET
480
C
A
440
PERFORM PRIORITY SCHEDULING ON THE QoS QUEUES IN THE SELECTED QoS CATEGORY
450
TIE ?
N
Y
460
PERFORM ROUND ROBIN OR LRU SCHEDULING ON THE QoS QUEUES IN THE SELECTED QoS CATEGORY
B

FIG.4

Attorney's Docket No.: 002717.P007

PATENT

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below, next to my name.

I believe I am the original, first, and sole inventor (if only one name is listed below) or an original, first, and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

POLICY BASED QUALITY OF SERVICE

the specification of which

_____ is attached hereto.

__X__ was filed on February 3, 1998 as United States Application Number 09/018,103 or PCT International Application Number _____ and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above.

I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d), of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s)			Priority Claimed	
(Number)	(Country)	(Day/Month/Year Filed)	Yes	No
(Number)	(Country)	(Day/Month/Year Filed)	Yes	No
(Number)	(Country)	(Day/Month/Year Filed)	Yes	No

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EL591666814 4US

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, DC 20231.

June 20, 2000
Date of Deposit

Heather S. South
Name of Person Mailing Correspondence

Signature 6/20/00
Date

I hereby claim the benefit under title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below

60/057371	8/29/97
(Application Number)	Filing Date
(Application Number)	Filing Date

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Number)	Filing Date	(Status -- patented, pending, abandoned)
(Application Number)	Filing Date	(Status -- patented, pending, abandoned)

I hereby appoint Aloysius T. C. AuYeung, Reg. No. 35,432; William Thomas Babbitt, Reg. No. 39,591; Jordan Michael Becker, Reg. No. 39,602; Bradley J. Bereznak, Reg. No. 33,474; Michael A. Bernadicou, Reg. No. 35,934; Roger W. Blakely, Jr., Reg. No. 25,831; Gregory D. Caldwell, Reg. No. 39,926; Kent M. Chen, Reg. No. 39,630; Lawrence M. Cho, Reg. No. 39,942; Thomas M. Coester, Reg. No. 39,637; Roland B. Cortes, Reg. No. 39,152; William Donald Davis, Reg. No. 38,428; Michael Anthony DeSanctis, Reg. No. 39,957; Daniel M. De Vos, Reg. No. 37,813; Tarek N. Fahmi, Reg. No. 41,402; James Y. Go, Reg. No. 40,621; Sharmini Nathan Green, Reg. No. 41,410; David R. Halvorson, Reg. No. 33,395; Eric Ho, Reg. No. 39,711; George W Hoover II, Reg. No. 32,992; Eric S. Hyman, Reg. No. 30,139; Dag H. Johansen, Reg. No. 36,172; Stephen L. King, Reg. No. 19,180; Michael J. Mallie, Reg. No. 36,591; Kimberley G. Nobles, Reg. No. 38,255; Ronald W. Reagin, Reg. No. 20,340; James H. Salter, Reg. No. 35,668; William W. Schaal, Reg. No. 39,018; James C. Scheller, Reg. No. 31,195; Charles E. Shemwell, Reg. No. 40,171; Maria McCormack Sobrino, Reg. No. 31,639; Stanley W. Sokoloff, Reg. No. 25,128; Allan T. Sponseller, Reg. No. 38,318; Steven R. Sponseller, Reg. No. 39,384; Judith A. Szepesi, Reg. No. 39,393; Edwin H. Taylor, Reg. No. 25,129; George G. C. Tseng, Reg. No. 41,355; Lester J. Vincent, Reg. No. 31,460; John Patrick Ward, Reg. No. 40,216; Ben J. Yorks, Reg. No. 33,609; and Norman Zafman, Reg. No. 26,250; my attorneys; and Robert Andrew Diehl, Reg. No. 40,992; Thomas A. Hassing, Reg. No. 36,159; and Edwin A. Sloane, Reg. No. 34,728; my patent agents, of BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP, with offices located at 12400 Wilshire Boulevard, 7th Floor, Los Angeles, California 90025, telephone (310) 207-3800; and James R. Thein, Reg. No. 31,710, my patent attorney; with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected herewith.

Send correspondence to Michael Anthony DeSanctis, BLAKELY, SOKOLOFF, TAYLOR &
(Name of Attorney or Agent)
ZAFMAN LLP, 12400 Wilshire Boulevard 7th Floor, Los Angeles, California 90025 and
direct telephone calls to Michael Anthony DeSanctis, (408) 720-8598.
(Name of Attorney or Agent)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Sole/First Inventor Stephen R. Haddock

Inventor's Signature _____ Date 4/24/98

Residence Los Gatos, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 18263 Bayview Drive
Los Gatos, CA 95030

Full Name of Second/Joint Inventor Justin N. Chueh

Inventor's Signature _____ Date 4/24/98

Residence Palo Alto, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 2333 Louis Road
Palo Alto, CA 94303

Full Name of Third/Joint Inventor Shehzad T. Merchant

Inventor's Signature _____ Date 4/24/98

Residence Mountain View, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 429 N. Rengstorff Avenue #1
Mountain View, CA 94043

Full Name of Fourth/Joint Inventor Andrew H. Smith

Inventor's Signature Andrew H. Smith Date 4/24/98

Residence Palo Alto, California (City, State) Citizenship United Kingdom (Country)

Post Office Address 336 Middlefield Road
Palo Alto, CA 94301

Full Name of Fifth/Joint Inventor Michael Yip

Inventor's Signature ______ Date 4/24/98

Residence Sunnyvale, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 1755 Chetamon Court
Sunnyvale, CA 94087

Title 37, Code of Federal Regulations, Section 1.56
Duty to Disclose Information Material to Patentability

(a) A patent by its very nature is affected with a public interest. The public interest is best served, and the most effective patent examination occurs when, at the time an application is being examined, the Office is aware of and evaluates the teachings of all information material to patentability. Each individual associated with the filing and prosecution of a patent application has a duty of candor and good faith in dealing with the Office, which includes a duty to disclose to the Office all information known to that individual to be material to patentability as defined in this section. The duty to disclosure information exists with respect to each pending claim until the claim is cancelled or withdrawn from consideration, or the application becomes abandoned. Information material to the patentability of a claim that is cancelled or withdrawn from consideration need not be submitted if the information is not material to the patentability of any claim remaining under consideration in the application. There is no duty to submit information which is not material to the patentability of any existing claim. The duty to disclosure all information known to be material to patentability is deemed to be satisfied if all information known to be material to patentability of any claim issued in a patent was cited by the Office or submitted to the Office in the manner prescribed by §§1.97(b)-(d) and 1.98. However, no patent will be granted on an application in connection with which fraud on the Office was practiced or attempted or the duty of disclosure was violated through bad faith or intentional misconduct. The Office encourages applicants to carefully examine:

(1) Prior art cited in search reports of a foreign patent office in a counterpart application, and

(2) The closest information over which individuals associated with the filing or prosecution of a patent application believe any pending claim patentably defines, to make sure that any material information contained therein is disclosed to the Office.

(b) Under this section, information is material to patentability when it is not cumulative to information already of record or being made or record in the application, and

(1) It establishes, by itself or in combination with other information, a prima facie case of unpatentability of a claim; or

(2) It refutes, or is inconsistent with, a position the applicant takes in:

(i) Opposing an argument of unpatentability relied on by the Office, or

(ii) Asserting an argument of patentability.

A prima facie case of unpatentability is established when the information compels a conclusion that a claim is unpatentable under the preponderance of evidence, burden-of-proof standard, giving each term in the claim its broadest reasonable construction consistent with the specification, and before any consideration is given to evidence which may be submitted in an attempt to establish a contrary conclusion of patentability.

(c) Individuals associated with the filing or prosecution of a patent application within the meaning of this section are:

(1) Each inventor named in the application;

(2) Each attorney or agent who prepares or prosecutes the application; and

(3) Every other person who is substantively involved in the preparation or prosecution of the application and who is associated with the inventor, with the assignee or with anyone to whom there is an obligation to assign the application.

(d) Individuals other than the attorney, agent or inventor may comply with this section by disclosing information to the attorney, agent, or inventor.